LA-UR-16-28155

Title: D-wave Quantum Computer as an Efficient Classical Sampler

Author(s): Chertkov, Michael
Hagberg, Aric Arild
Lokhov, Andrey
Misiakiewicz, Theodor
Misra, Sidhant
Vuffray, Marc Denis

Intended for: Distribution of slides from the ISTI D-wave rapid response call

Issued: 2016-10-26 (Draft)

D-wave Quantum Computer as an Efficient Classical Sampler

M. Chertkov[1,2] (PI), A. Hagberg[3], A. Lokhov[1,2] (co-PI),
T. Misiakiewicz[1], S. Misra[3], M. Vuffray[2]

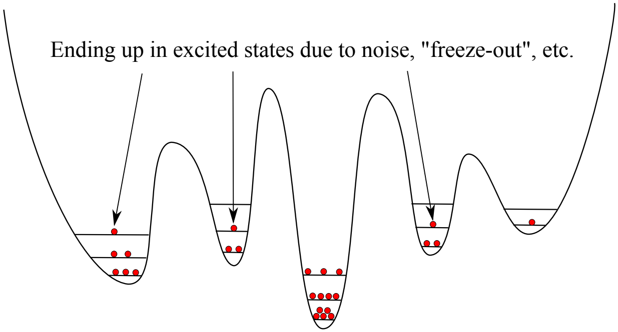
[1]Center for Nonlinear Studies
[2]Theoretical Division T-4
[3]Theoretical Division T-5

D-wave Quantum Computing Efforts Debrief

Introduction: D-wave as an efficient sampler

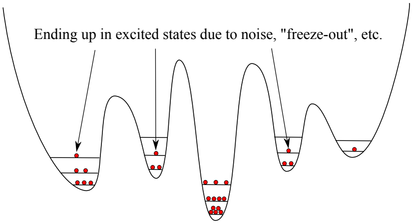
Theoretical and experimental evidence that D-wave can approximately sample from a Boltzmann distribution at some effective temperature

Ronnow *et al.*, Science (2014)
Amin, Phys. Rev. A (2015)
Perdomo-Ortiz *et al.*, Sci. Rep. (2016)
Benedetti *et al.*, Phys. Rev. A (2016)



Introduction: D-wave as an efficient sampler

Disadvantage for optimization turned into advantage for numerous applications:



- ✓ Restricted Boltzmann Machines (blocks for Deep Learning)

  Denil & Freitas, NIPS (2011); Dumoulin *et al.*, AAAI Artificial Intelligence (2015); Benedetti *et al.*, Phys. Rev. A (2016); Amin *et al.*, "Quantum Boltzmann Machine" (2016)

- ✓ Producting samples in hard glassy models

  Katzgraber *et al.*, Phys. Rev. X (2014 & 2015); Martin-Mayor & Hen, Sci. Rep. (2015); Venturelli *et al.*, Phys. Rev. X (2015); Zhu *et al.*, Phys. Rev. A (2016)

- ✓ Accurate calibration of the D-wave machine

  King & McGeoch (2014) "Algorithm engineering for a quantum annealing platform"; Perdomo-Ortiz *et al.*, Sci. Rep. (2016); Raymond *et al.*, "Global warming: temperature estimation in annealers" (2016); Also example in this debrief!

Relation between input and effective Hamiltonians in D-wave

Input Hamiltonian

$$\mathcal{H} = \sum_{\langle i,j \rangle} J_{ij}\sigma_i\sigma_j + \sum_{i \in V} H_i\sigma_i$$

Effective Hamiltonian in D-wave

$$\mathcal{H}_{\text{eff}} = \sum_{\langle i,j \rangle} J'_{ij}\sigma_i\sigma_j + \sum_{i \in V} H'_i\sigma_i$$

Relation between input and effective Hamiltonians in D-wave

Input Hamiltonian

$$\mathcal{H} = \sum_{\langle i,j \rangle} J_{ij}\sigma_i\sigma_j + \sum_{i \in V} H_i\sigma_i$$

Effective Hamiltonian in D-wave

$$\mathcal{H}_{\text{eff}} = \sum_{\langle i,j \rangle} J'_{ij}\sigma_i\sigma_j + \sum_{i \in V} H'_i\sigma_i$$

Let us write $J'_{ij} = \beta(J_{ij} + \Delta J_{ij})$, $H'_i = \beta(H_i + \Delta H_i)$, where

$T = 1/\beta$: effective temperature

ΔJ_{ij}, ΔH_i : possible biases

Correspondence $\mathcal{H} \leftrightarrow \mathcal{H}_{\text{eff}}$ by solving the reconstruction problem of learning β, ΔJ_{ij}, ΔH_i from samples produced by D-wave with $\mathcal{H}_{\text{eff}}$

Reconstruction problem in D-wave

Given M independent samples (configurations), reconstruct $\mathcal{H}_{\text{eff}}$

	$k = 1$	$k = 2$	...	$k = \mathrm{M}$
σ_1	$+1$	-1	...	$+1$
σ_2	-1	-1	...	-1
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
σ_{N}	$+1$	$+1$	...	-1

Task known as **Inverse Ising Problem**. The optimal algorithm for solving this task is the LANL-developed "Screening method"

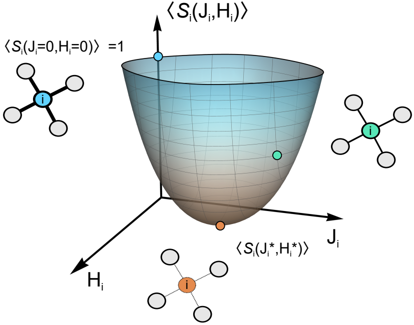
Vuffray, Misra, Lokhov, Chertkov, NIPS (2016)
Lokhov, Vuffray, Misra, Chertkov, submitted to Nature Physics (2016)

How does Screening method work?

For each spin, minimize the potential $S_i(J_i, H_i)$ which applies counter-interactions ($P \propto e^{-\mathcal{H}}$):

$$(\widehat{J}_i, \widehat{H}_i) = \underset{(J_i, H_i)}{\operatorname{argmin}} \Big(S_i(J_i, H_i) + \lambda \|J_i\|_1 \Big)$$

$$S_i(J_i, H_i) = \langle \exp(\sum_{j \neq i} J_{ij}\sigma_i\sigma_j + H_i\sigma_i) \rangle_M$$



Vuffray, Misra, Lokhov, Chertkov, NIPS (2016)
Lokhov, Vuffray, Misra, Chertkov, submitted to Nature Physics (2016)

First outcome of this project: development of an efficient algorithmic implementation using advanced first-order optimization methods ($\sim N^2$ times faster, to appear on GitHub)

Effective temperature

Where does the effective temperature come from? Let us look at the annealing procedure with $\tau = t/t_{\text{annealing}}$:

$$\mathcal{H}(\tau) = A(\tau)\left(-\sum_{i\in V}\sigma_i^x\right) + B(\tau)\left(\sum_{\langle i,j\rangle} J_{ij}\sigma_i^z\sigma_j^z + \sum_{i\in V} H_i\sigma_i^z\right)$$

Monotonic functions A and B satisfy $A(0) \gg B(0)$ and $A(1) \ll B(1)$.

Effective temperature

Where does the effective temperature come from? Let us look at the annealing procedure with $\tau = t/t_{\text{annealing}}$:

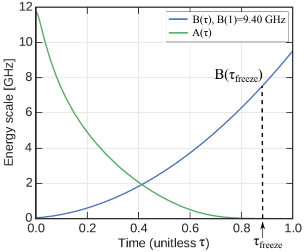
$$\mathcal{H}(\tau) = A(\tau)\left(-\sum_{i\in V}\sigma_i^x\right) + B(\tau)\left(\sum_{\langle i,j\rangle} J_{ij}\sigma_i^z\sigma_j^z + \sum_{i\in V} H_i\sigma_i^z\right)$$

Monotonic functions A and B satisfy $A(0) \gg B(0)$ and $A(1) \ll B(1)$.

The "freeze-out" phenomenon: the evolution stops at the point τ_{freeze}:

$$T_{\text{eff}} = T_{\text{D-wave}}\frac{B(1)}{B(\tau_{\text{freeze}})}$$

Benedetti *et al.*, Phys. Rev. A (2016)
Raymond *et al.*, "Global warming: temperature estimation in annealers" (2016)

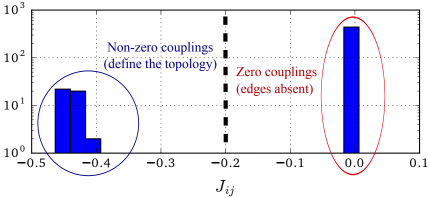


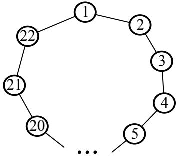
✓ **No unique T_{eff}:** β is the function of the input Hamiltonian

✓ **"Single qubit freeze-out":** τ_{freeze} can vary for different spins

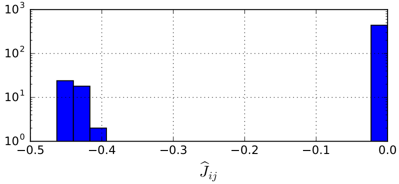
Illustration: estimating the effective temperature

Data set (from **Marcus Daniels**): embedded closed circles of $N = 22$ spins with different values of $J_{i,i+1}$ and $H_i = 0$ (diverse realizations, $t_{\text{annealing}}$, *etc.*). Example for $M = 7250$ and $J_{i,i+1} = -0.0625\ \forall(i, i+1)$.





Refined $\{J'_{ij}, H'_i\}$. Neglecting H'_i and biases, $\beta_{\text{eff}} \approx 7$ since $\overline{J'} = -0.44$.



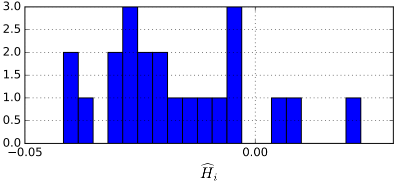
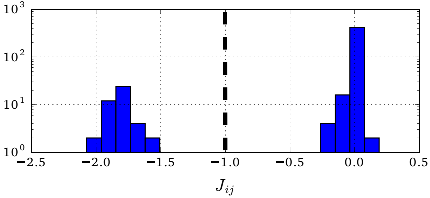
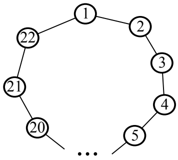


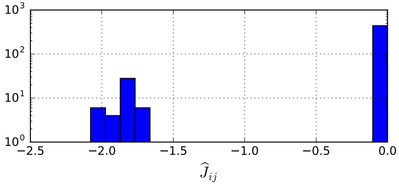
Illustration: estimating the effective temperature

Example for $M = 7250$ and $J_{i,i+1} = -0.4375\ \forall(i, i+1)$.





Refined $\{J'_{ij}, H'_i\}$. Neglecting H'_i and biases, $\beta_{\text{eff}} \approx 4.2$ since $\overline{J'} = -1.84$.



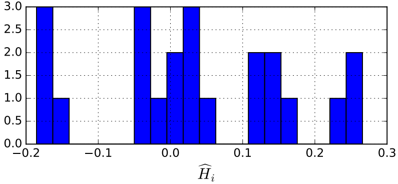
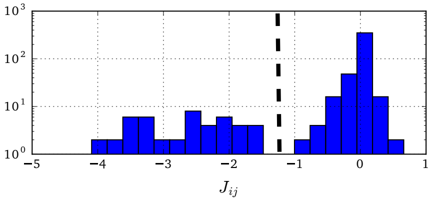
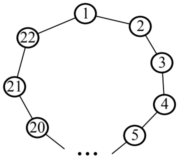


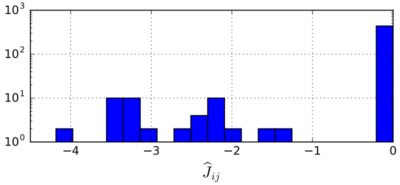
Illustration: estimating the effective temperature

Example for $M = 7250$ and $J_{i,i+1} = -0.75\ \forall (i, i+1)$.





Refined $\{J'_{ij}, H'_i\}$. Neglecting H'_i and biases, $\beta_{\text{eff}} \approx 3.72$ since $\overline{J'} = -2.79$.



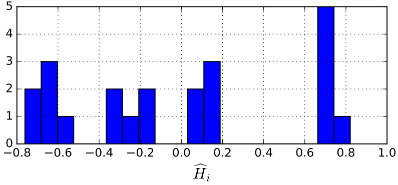
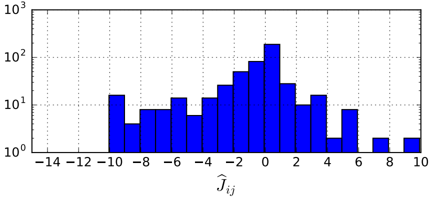
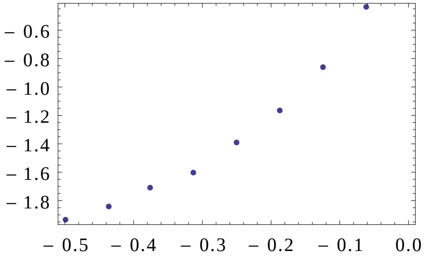


Illustration: estimating the effective temperature

In the case of $J_{i,i+1} = -1.0\ \forall (i, i+1)$, $M = 7250$ is insufficient: the topology can not be correctly recovered.



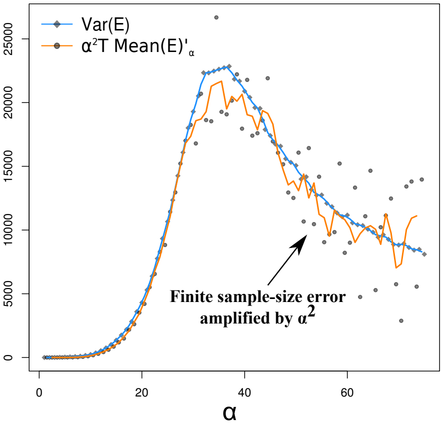
Dependence between J' on J:



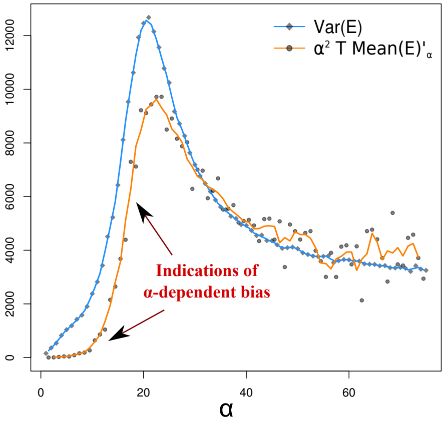
What about biases?

Simple test: if $P(\underline{\sigma}) \propto e^{-\mathcal{H}(\underline{\sigma})/(\alpha T)}$, then $\alpha^2 T \frac{\partial}{\partial \alpha}\langle H \rangle = \langle H^2 \rangle - \langle H \rangle^2$

Checkerboard pattern with magnetic fields



Checkerboard pattern without magnetic fields

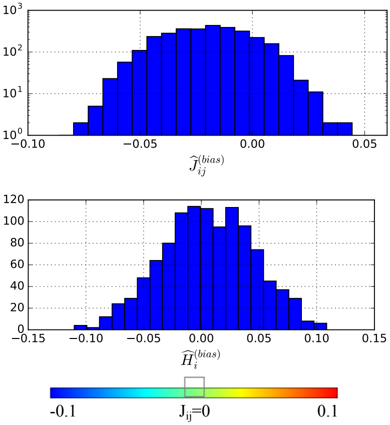


Example found by **Carleton Coffrin**, see next talk!

Illustration: detecting and correcting biases

Example of the input $\mathcal{H} = 0$ over the entire Chimera graph

Although D-wave comes with a software for correcting biases, they are still present and persist. Example from the Burnaby machine on Sep 15:



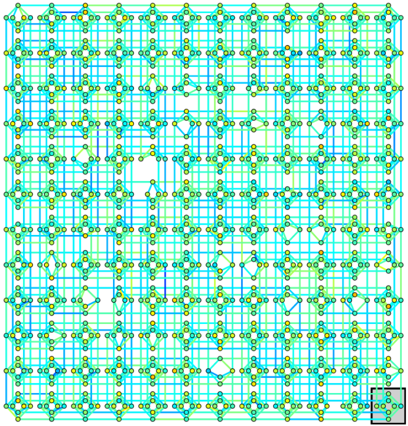


Illustration: detecting and correcting biases

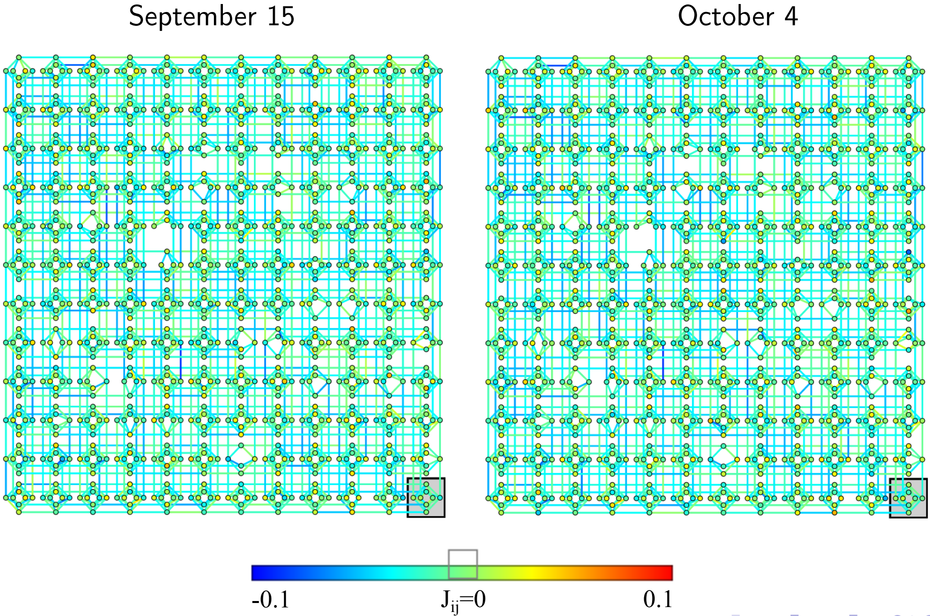
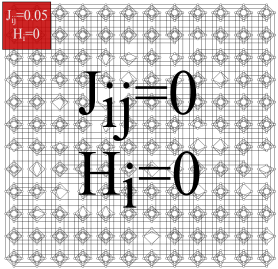
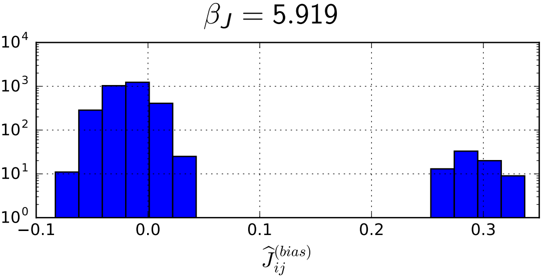
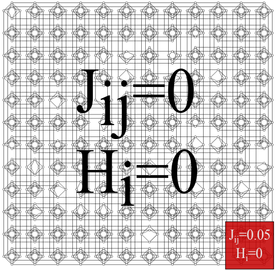


Illustration: detecting and correcting biases







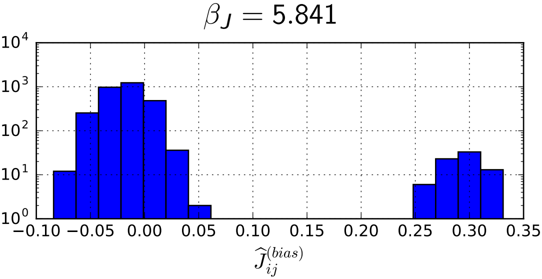
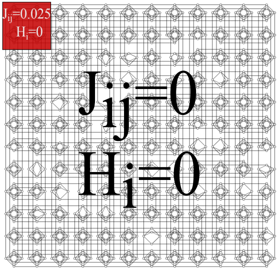
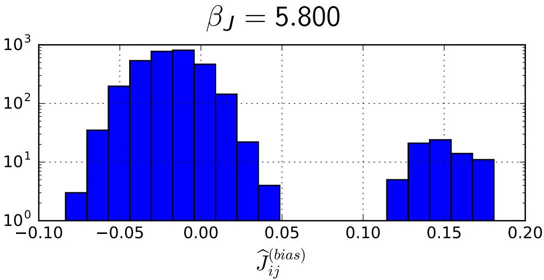
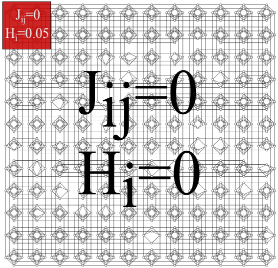


Illustration: detecting and correcting biases







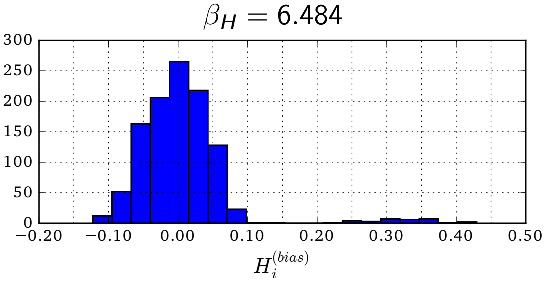
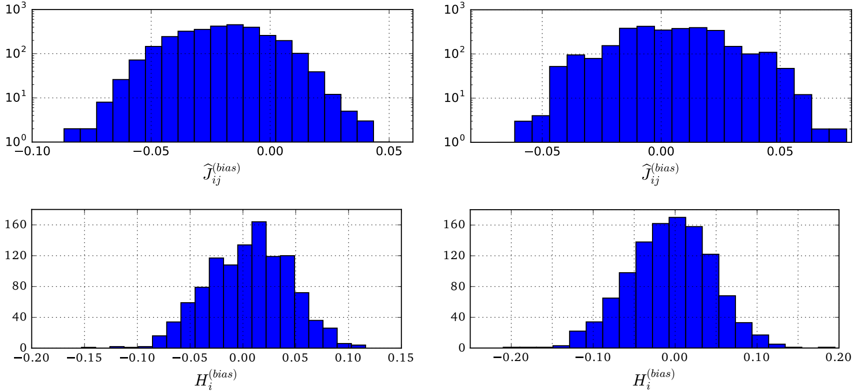


Illustration: detecting and correcting biases

Corrections: inputting $\mathcal{H} = -\frac{1}{\beta_J}\sum_{\langle i,j\rangle} J_{ij}^{(bias)}\sigma_i\sigma_j - \frac{1}{\beta_H}\sum_{i\in V} H_i^{(bias)}\sigma_i$

$\mathcal{H}_{\text{eff}}$ before corrections $\mathcal{H}_{\text{eff}}$ after corrections



Symmetrized and more squeezed distributions with a single iteration

Path forward: efficient calibration of the D-wave machine

The calibration issue addressed in several recent papers with heuristic methods: King & McGeoch (2014); Perdomo-Ortiz *et al.*, Sci. Rep. (2016); ...

As shown in the previous examples, **we can do much better!**

✓ Iteratively correcting the biases for the target $\mathcal{H}_T$:

$$
\begin{aligned}
&i)\ \frac{\mathcal{H}_T}{\beta} \longrightarrow \mathcal{H}_T + \Delta(\mathcal{H}_T) \\
&ii)\ \frac{\mathcal{H}_T - \Delta(\mathcal{H}_T)}{\beta} \longrightarrow \mathcal{H}_T - \Delta(\mathcal{H}_T) + \Delta(\mathcal{H}_T - \Delta(\mathcal{H}_T)) \\
&\qquad\qquad\qquad\qquad \approx \mathcal{H}_T - \Delta'(\mathcal{H}_T)\Delta(\mathcal{H}_T) \\
&iii)\ \frac{\mathcal{H}_T - \Delta(\mathcal{H}_T) + \Delta'(\mathcal{H}_T)\Delta(\mathcal{H}_T)}{\beta} \longrightarrow \ldots
\end{aligned}
$$

✓ **Machine learning task:** learn the functional form of $\Delta(\mathcal{H}_T)$ with the linear response theory; start directly at the point (*ii*)

✓ Include the **higher-order interaction** terms in the reconstruction problem to capture the effect of inactive spins

Acknowledgements and questions

Many thanks to **Marcus Daniels** for the data set used in the first part of the work, and to **Carleton Coffrin** for the insight and data contributions in the second part!

Questions?